Hourly energy demand and prices: an analysis on risk measures and correlation

Camilo Oberndorfer Mejía   Gregorio Pérez Bernal   Luisa Toro Villegas

Miguel Valencia Ochoa

May 24, 2021

Abstract

blah blah lo que vamos a usar blah blah.

The previously mentioned methodologies will be applied to the hourly energy demand generation in Spain from 2015 to 2019. The hope to this paper is to find a way to predict energy prices using correlation techniques and understand how demand affects supply and vicersa.

1 Introduction

2 Data description

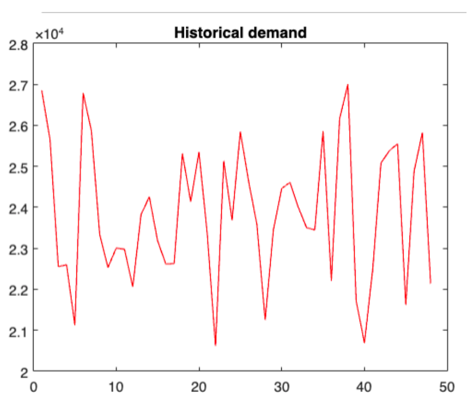
2.1 Original dataset

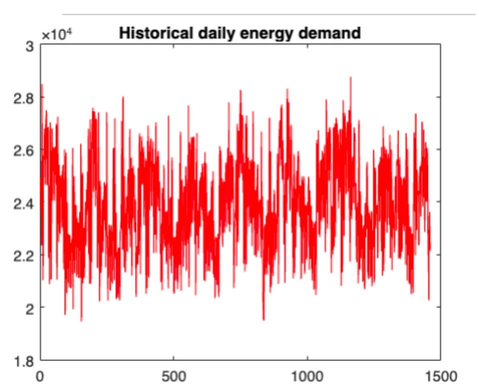
2.2 Modifications

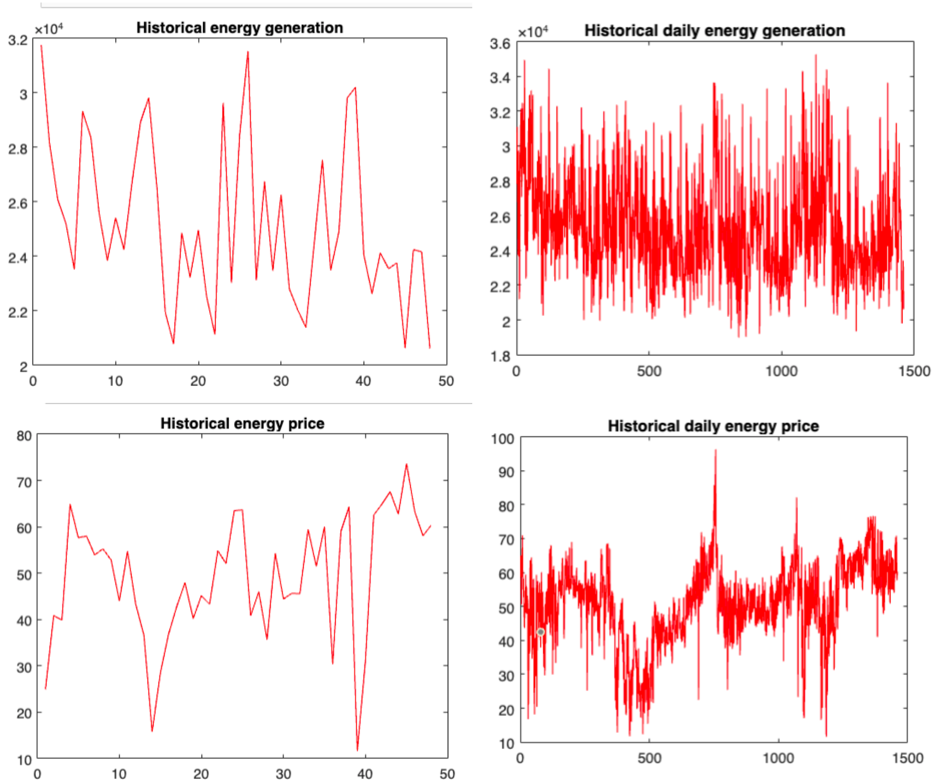
Montly= Demand

3 Estimation and results

Historical data



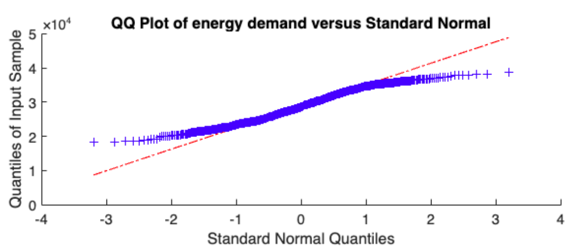


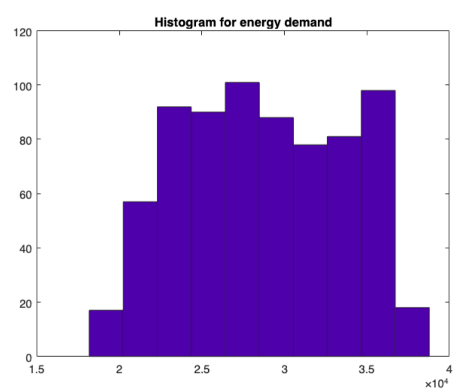


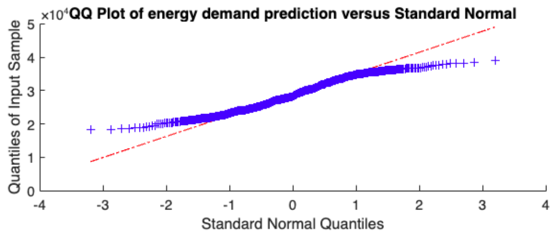
3.1 Identifying underlying distribution

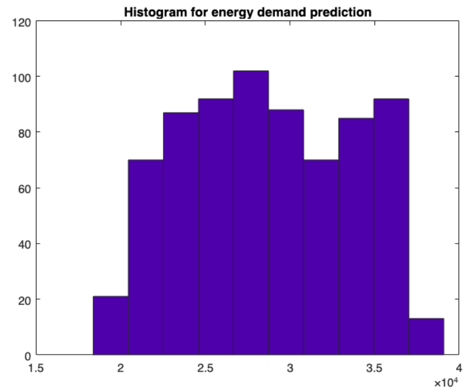
To construct a thorough analysis on any given set of data, a crucial first step is to attempt to understand that way the data is distributed. If the data follow a normal distribution, the following analysis could have a smaller margin of error, but it is rare to see that any real life data follows a normal distribution.

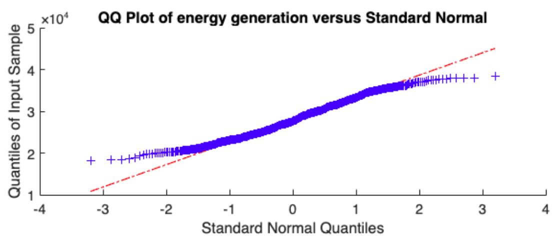
The data's energy generation histogram and quantile-quantile plot (qqplot) are graphic tools that help identify normality, which are shown in the following figures:

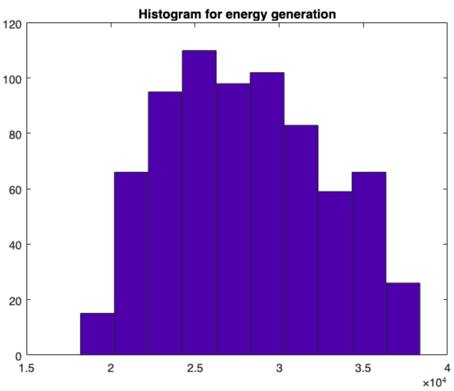












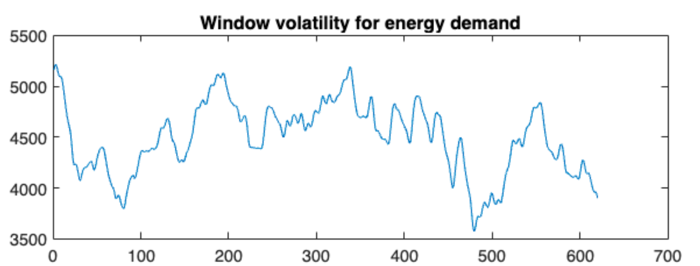
The previous figures are a characteristic example of under-dispersed data (for example a uniform distribution or any other beta distribution) and the qqplot, which compares the theoric quantiles and sample quantiles, shows an s-shape, which bleah blah blah.

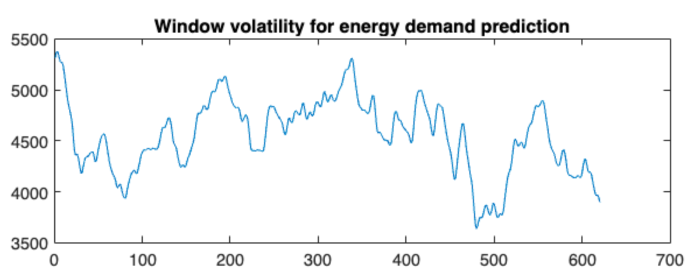
Linealizacion?

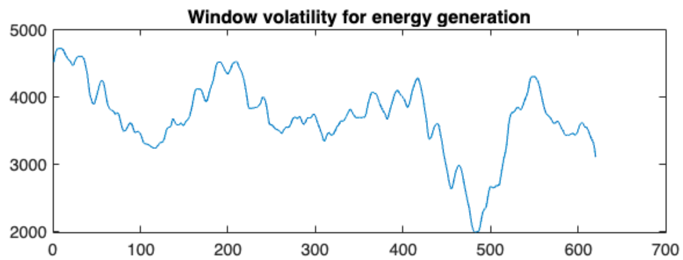
3.2 Stability tests

Given the data's underlying bleh bleh

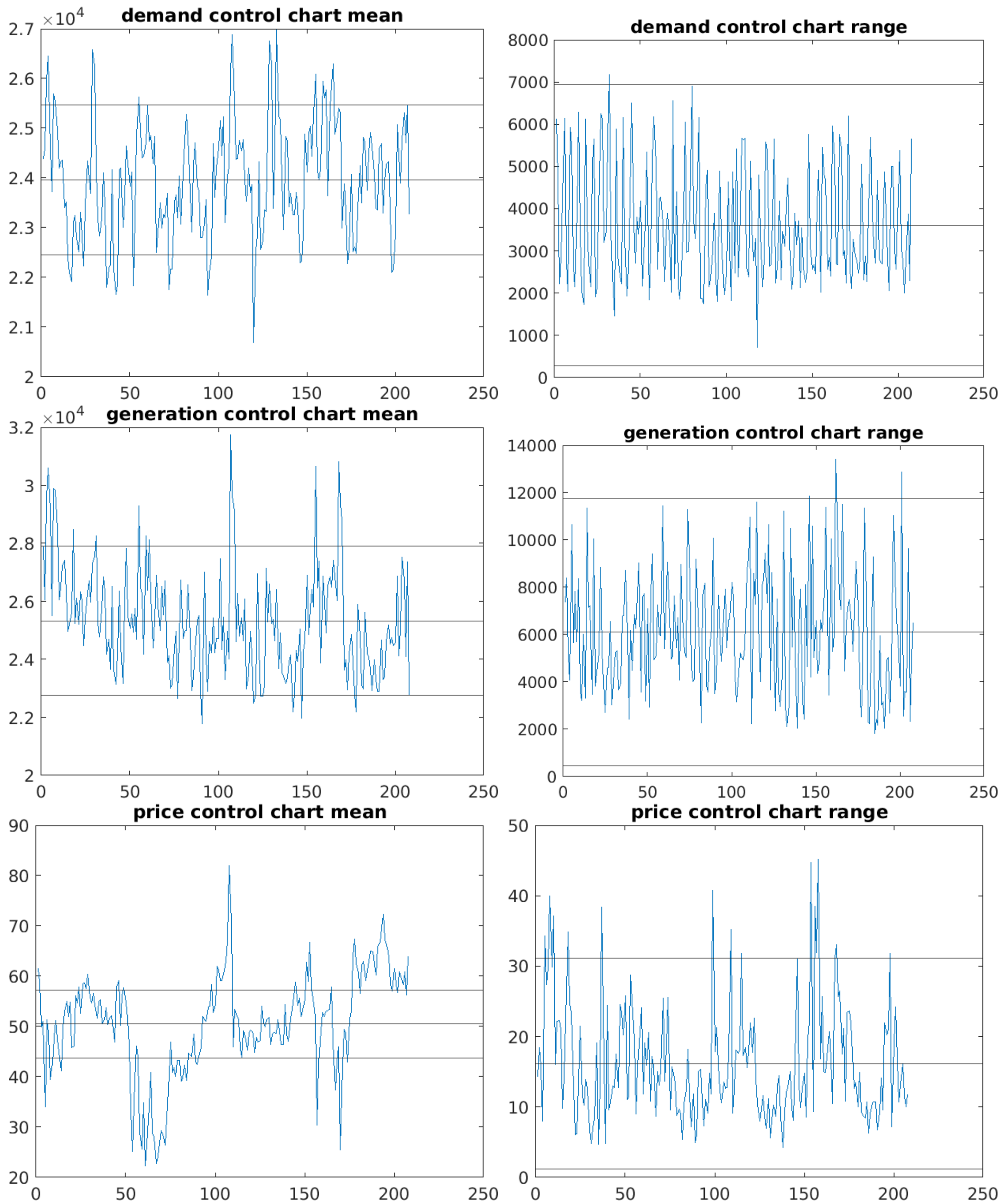
window volatility (ya) and asymptotic volatility(si alguien lo encuentra)





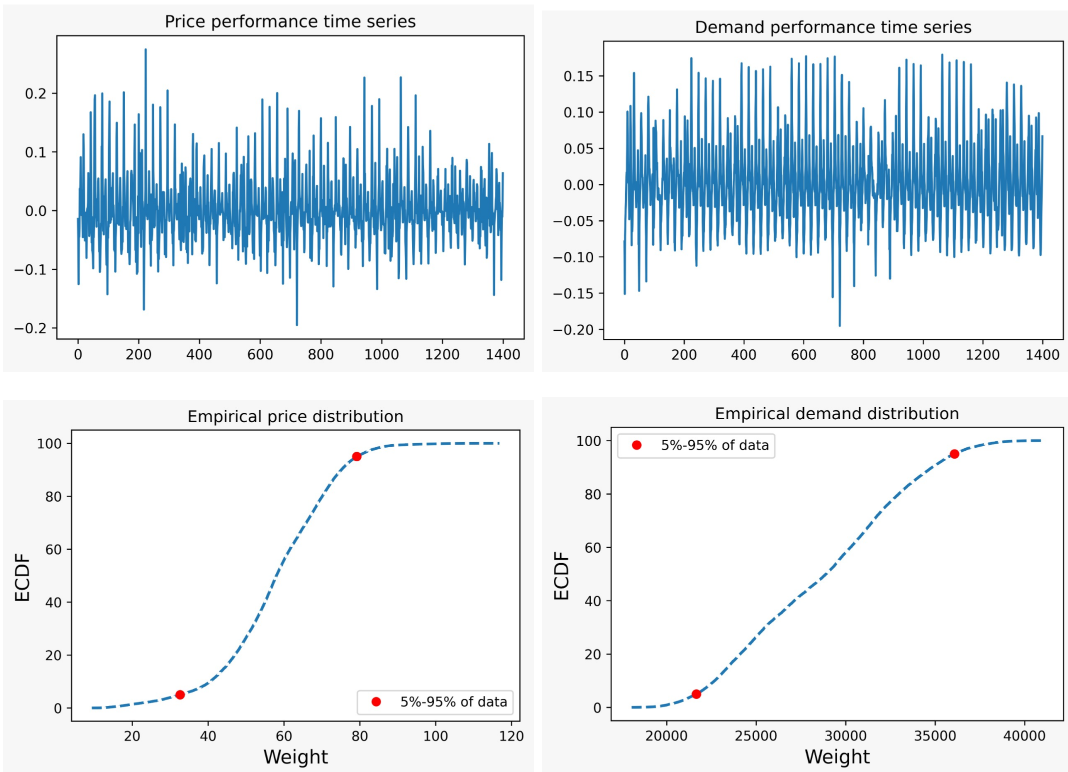


Cartas de control: hablar sobre esto.



Estimaciones de las volatilidades con el método de suavizamiento exponencial

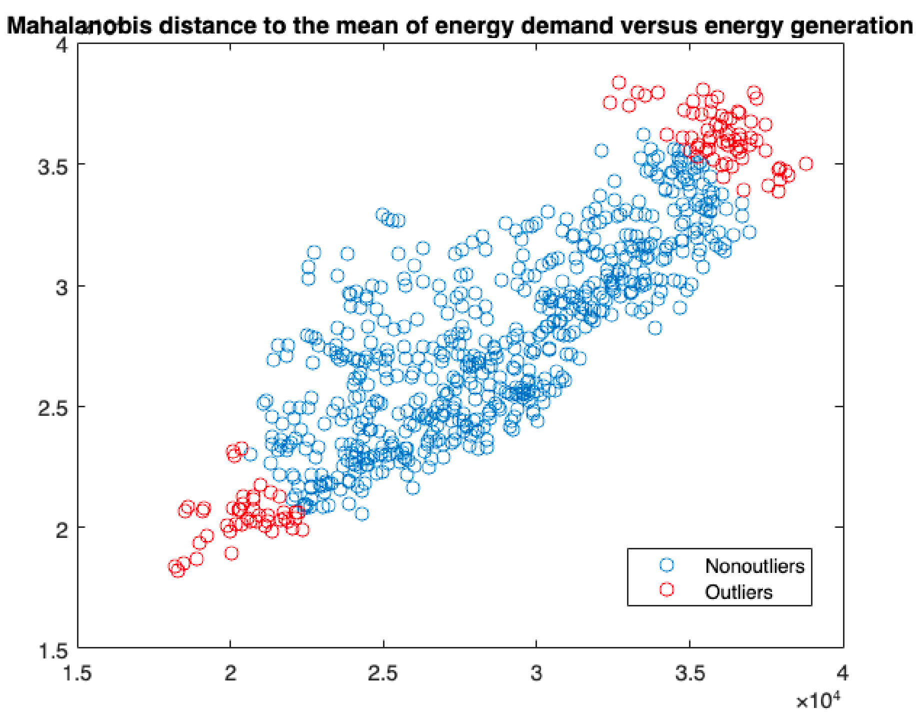
3.3 Performance and prediction error



3.4 Extreme values and outliers

Metricas (Distancia euclidiana, mahalanobis)

Mahalanobis distance is used to detect outliers in multivariant testing.



¿Cuales otras comparaciones hacemos?

EVT ? El método de picos sobre el umbral, POT por sus siglas en inglés, identifica los valores extremos de la serie de retornos como aquellos que excedan un umbral u , estos valores son conocidos como excesos de retorno39

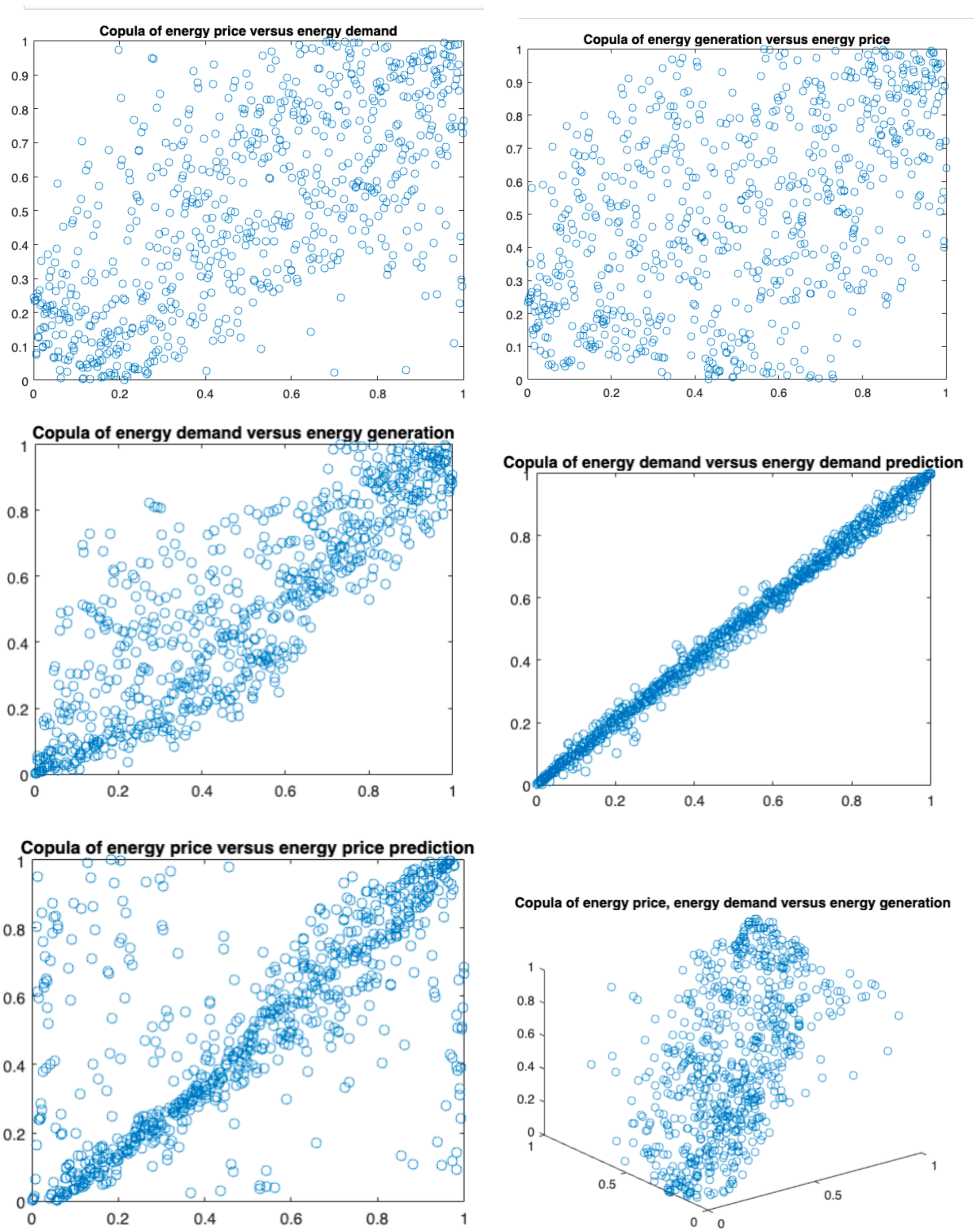
Funciones de densidad (izquierda) y distribución (derecha) de las distribuciones de valor extremo.

partir de la teoría del valor extremo es posible introducir otras medidas de riesgo que se concentran en la frecuencia y la magnitud de la realización de eventos extremos. Estas medidas son conocidas como el Return level y el Return period.

3.5 Measuring dependency

las funciones de autocorrelación simple y parcial, FAC y FACP, (hay que buscar que es eso)

A copula helps detect dependency structures in multivariate data.



Coeficiente de correlacion (R2)

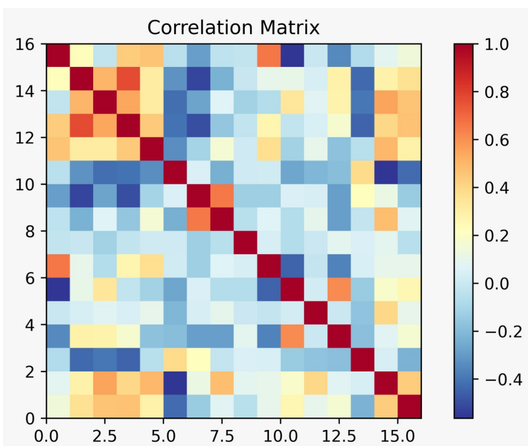
Table	TotalEnergyGeneration	EnergyDemand	EnergyPrice
"Correlation coeficient between variables (R2)"	0.74797	0.81922	0.46958

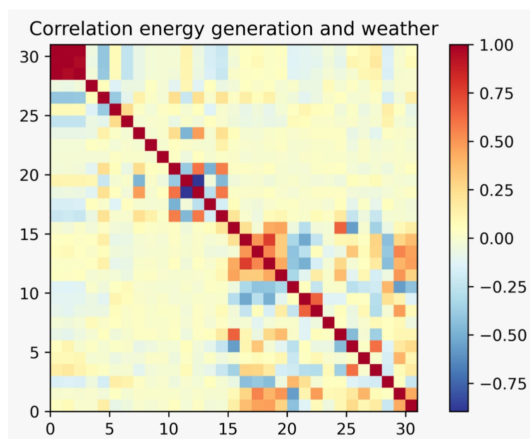
Coeficiente de correlacion (R2) con matriz de covarianzas robusta

Table	TotalEnergyGeneration	EnergyDemand	EnergyPrice
"Correlation coeficient between variables (R2)"	0.82075	0.83201	0.27476

Coeficiente de correlacion (R2) con matriz de covarianzas robusta entre dos variables de interés

Table	TotalEnergyGeneration	EnergyDemandPrediction
"Correlation coeficient between variables (R2)"	0.83751	0.83751





proecciones
Esto no se puede borrar, hay que buscar donde ponerlo [MV06] [GP09] [Jha19]

References

[GP09] Humberto Gutiérrez Pulido. Control estadístico de la calidad y seis sigma. Segunda Edición, 2009.

[Jha19] Nicolas Jhana. Hourly energy demand generation and weather. *Kaggle*, 2019.

[MV06] Luis Fernando Melo Velandia. Medidas de riesgo, características y técnicas de medición: una aplicación del var y el es a la tasa interbancaria de colombia. 2006.